

An Extremely Accurate MaxT Forecast Using the Analog Tool in GFE

**Les Colin
WFO Boise, ID**

I. Introduction

The GFE grid-archiving and verification framework known as BOIVerify (Barker, 2006) has made possible the development of two powerful methods for forecasting daily Max and Min temperatures, and morning and afternoon dew points (TdMrn and TdAft). One of these methods is the bias-corrected system (Barker, 2006), now widely used throughout the Western U.S. The other is the Analog tool, developed by Les Colin and Steve Parker at WFO Boise in 2007 (Colin, 2008).

This article will trace how the Analog tool produced MaxT grids for the BOI CWA for July 13, 2008 that verified as the most accurate forecast temperature grids ever made at WFO BOI.

II. A Brief Description of the Analog Tool

The Analog tool consults archived grids within BOIVerify to find pattern matches for a current model forecast grid. BOIVerify retrieves all past grids for the given model, date, and forecast period, along with the observed grids for those same dates. The pattern matching algorithm is fully explained in the documentation for the Analog14 tool, available on the STR. The Analog tool allows examination of up to 50 past days of archived data from which only a few (the ones most similar to the current forecast grid) are actually used. A typical number for days-used is 12. The tool then retrieves the observed grids for those same dates and from these computes error-grids made by the archived forecasts. These error-grids are weighted according to how similar each archived forecast grid was to the current forecast grid. From this a weighted average error-grid is computed and then applied as a correction to the current grid. The Analog tool allows blending of many different models, varying the number of days examined and days-used in the archives, and selection of only those archived dates which were common to every model in a given blend.

III. July 13, 2008

The blend used for MaxT on July 13 consisted of 5 models: GFS40, NAM12, ADJMAV, ADJMET, and MosGuide, with 40 past days examined and 12 used. This is a blend commonly used at WFO BOI. The common-dates option was turned off.

Figure 1 shows the original first period GFS40 grid for July 13 as it was smart-initiated onto the BOI domain. If the forecaster had used this grid as the actual forecast it would have verified according to the histogram in Figure 2. Figure 3 shows the error-grid made by the GFS40.

But if the 40/12 Analog choice had been used for that GFS40, it would have produced the weighted average error-grid shown in Figure 4. If that weighted error-grid had been used to correct the original GFS40 grid the result would have verified as shown in Figure 5, a huge improvement. The errors are shown in Figure 6. In fact, if only the single best analog had been used in the 40 days prior to July 13 (the one from July 9, 2008) the correction made by that error-grid alone would have produced the histogram in Figure 7 and the errors in Figure 8.

The other models in the blend made similar improvements, but not always in the same places. Errors made by 40/12 choices for the other models in the blend (of five models) taken separately are shown in Figures 9-12.

When all those model corrections were combined the result was even better than for any of the models taken separately. The BOI CWA contains 22199 “points” (where the “points” are really 2.5 km square boxes). The histogram for the blend scored 99.40 on all the points in the BOI CWA, or even better than for any individual model in the blend. (The score is defined as 100 minus the mean squared error for all points. A score of 99.40 means a mean squared error of 0.60 or a mean absolute error of 0.77F degrees.) Evidently the models were correcting errors made by each other. Figure 13 shows the errors for the blend, Figure 14 shows the histogram, and Figure 15 shows the final 40/12 Analog forecast blend alongside the observed MaxT grid for July 13.

Incidentally, the histogram for the BC blend using the same models scored 99.02. No original forecast from any model, nor from any original model blend, has ever come close to these results.

The Analog system produced MaxT histogram scores of 99+ as far out as period 4. Before July 13, 2008, the highest score ever achieved by a human forecaster at WFO BOI had been 99.06 on a first-period MaxT valid Sep 29, 2006.

IV. Discussion

After collecting Analog histogram scores in periods 1-14 for about two years, we found that the very best MaxT verification scores have tended to occur during the warm season in the warming days following passage of a cold front. The case of July 13, 2008 fit that pattern well, as a relatively strong cold front (for July) passed through around July 10. This is an empirical finding only—the meteorological causes are not well understood, but at least we found a recognizable pattern.

We also found that Analog-derived TdMrn and TdAft have tended to do better than BC-derived TdMrn and TdAft, but again the meteorological reasons are unknown.

Finally, we found that Analog grids have tended to do slightly better than BC grids in the early forecast periods (1-4), but beyond period 7 the BC grids have tended to do better. Perhaps this is because in later periods the archived grids differ so much from the current grids and from each other that the corrections suggested by even the best Analogs are increasingly random. The BC method doesn't seem to have that problem.

V. References

Barker, T.W., 2006: BOIVerify 1.0,
<http://www.mdl.nws.noaa.gov/~applications/STR/generalapplinfoout.php3?appnum=1089>

Colin, L.R., 2008: Analog14 1.4,
<http://www.mdl.nws.noaa.gov/~applications/STR/generalapplinfoout.php3?appnum=1223>

VI. Figures

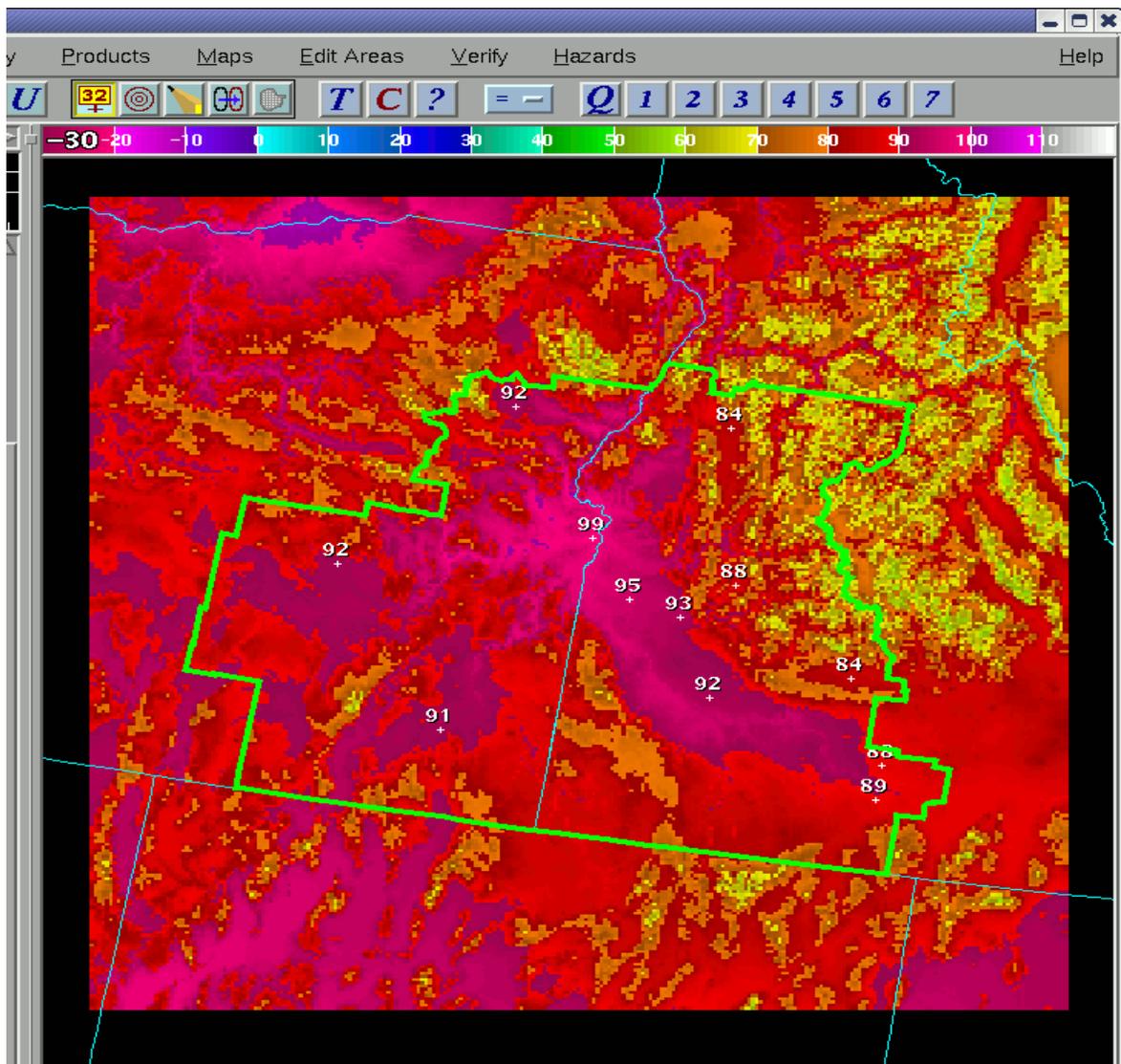


Fig 1. Original GFS40 first period forecast of MaxT for July 13, 2008. The BOI CWA, which covers southeastern Oregon and southwestern Idaho, is outlined in green.

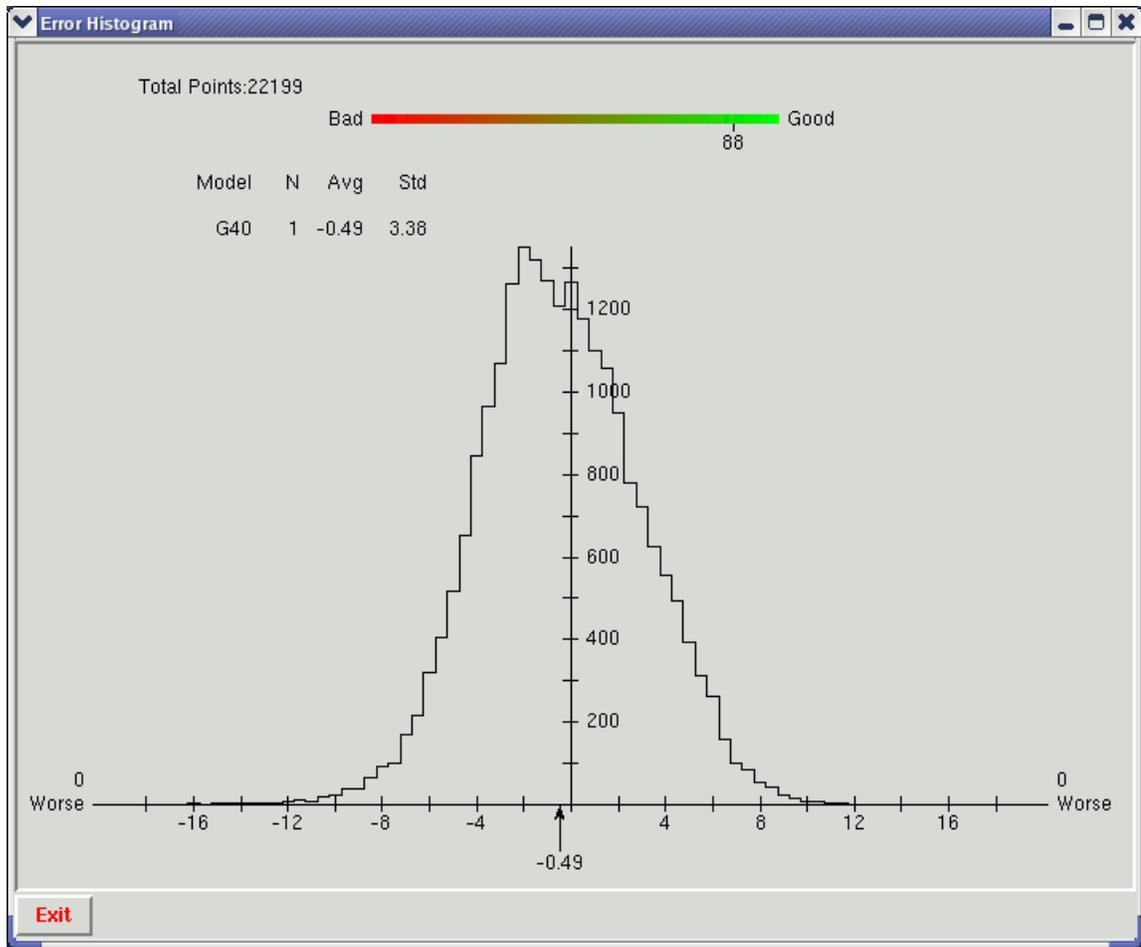


Fig 2. Error histogram for GFS40 forecast of Fig 1. To two decimal places this histogram scored 88.34. Note the wide histogram spread indicating numerous errors greater than $\pm 3F$ degrees.

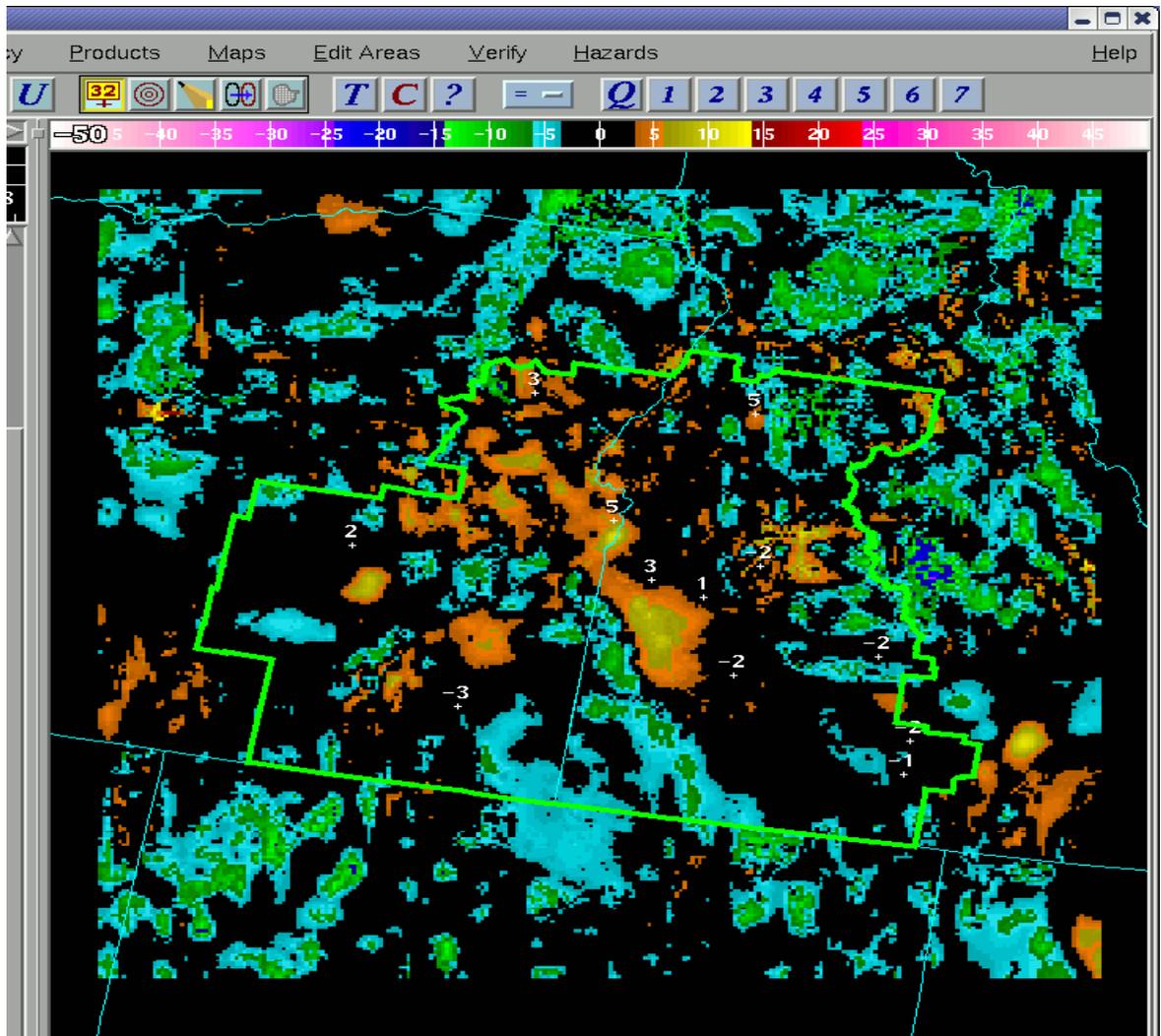


Fig 3. GFS40 MaxT error grid for Fig 1. Colored areas show where the errors were greater than $\pm 3F$ degrees.

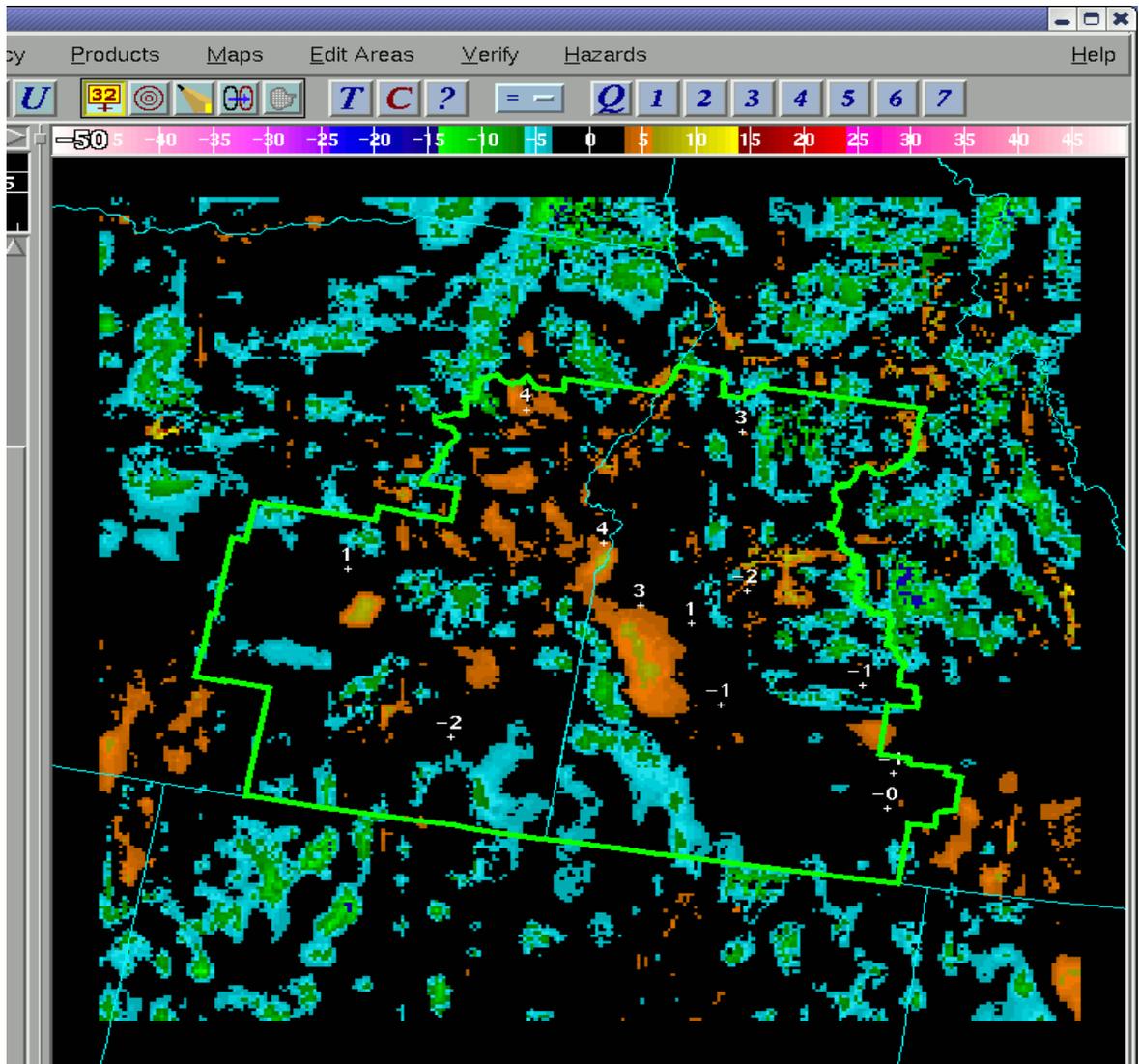


Fig 4. MaxT weighted error correction grid obtained from 40/12 Analog choice on the GFS40 of Fig 1.

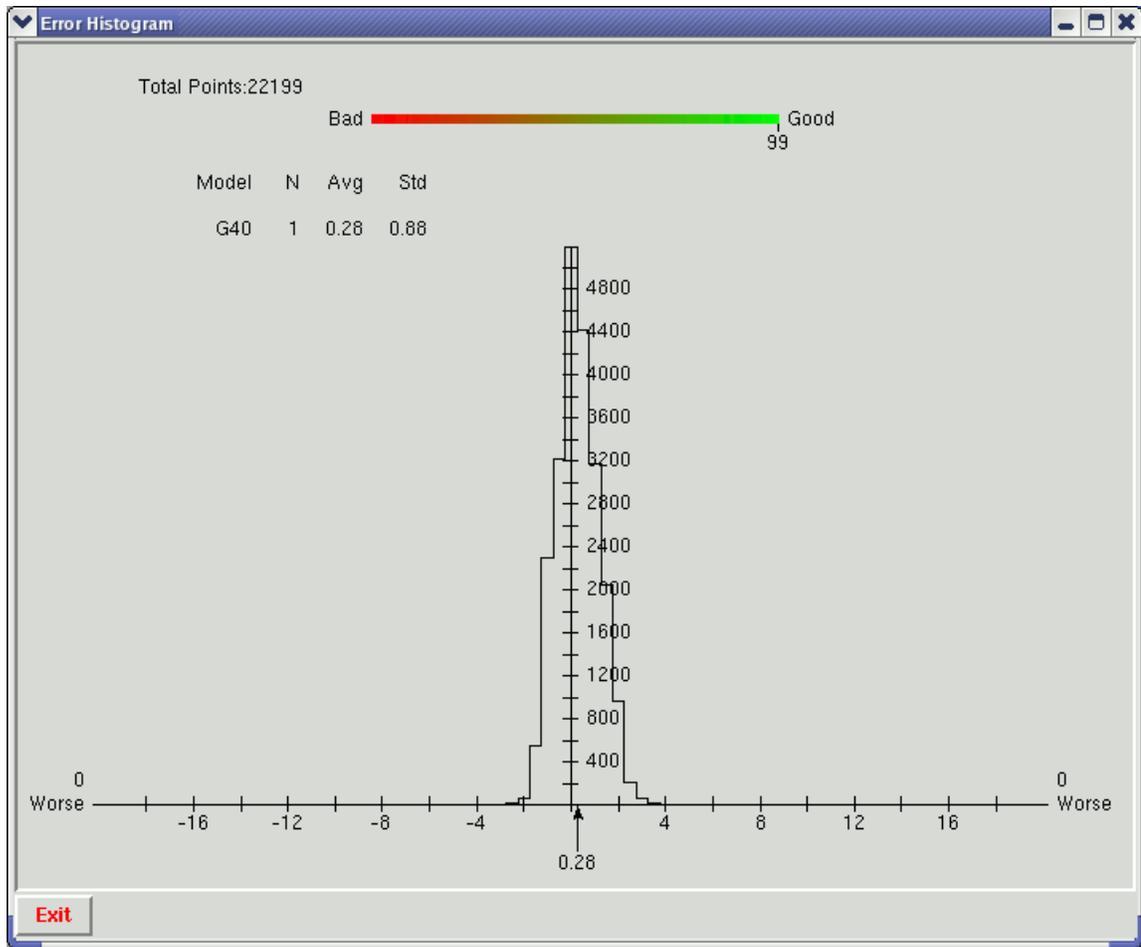


Fig. 5 Error histogram after applying the error-grid in Fig 4 to the original GFS40 in Fig 1. This histogram scored 99.14 and was (temporarily) the highest scoring histogram for MaxT ever produced at WFO BOI.

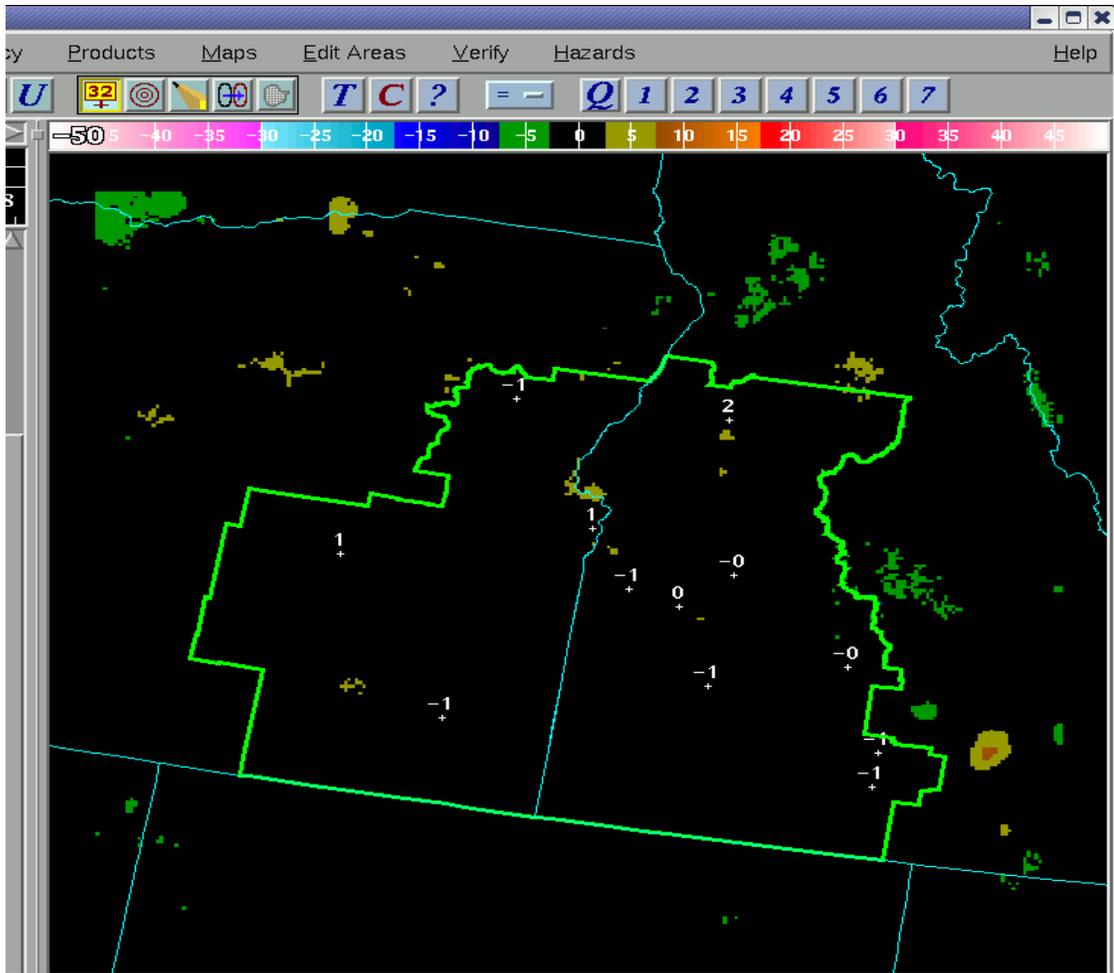


Fig. 6. MaxT errors after applying the correction grid of Fig 4 to the original GFS40 in Fig 1. Note that nearly every error within the CWA (green boundary) was within $\pm 3F$ degrees.

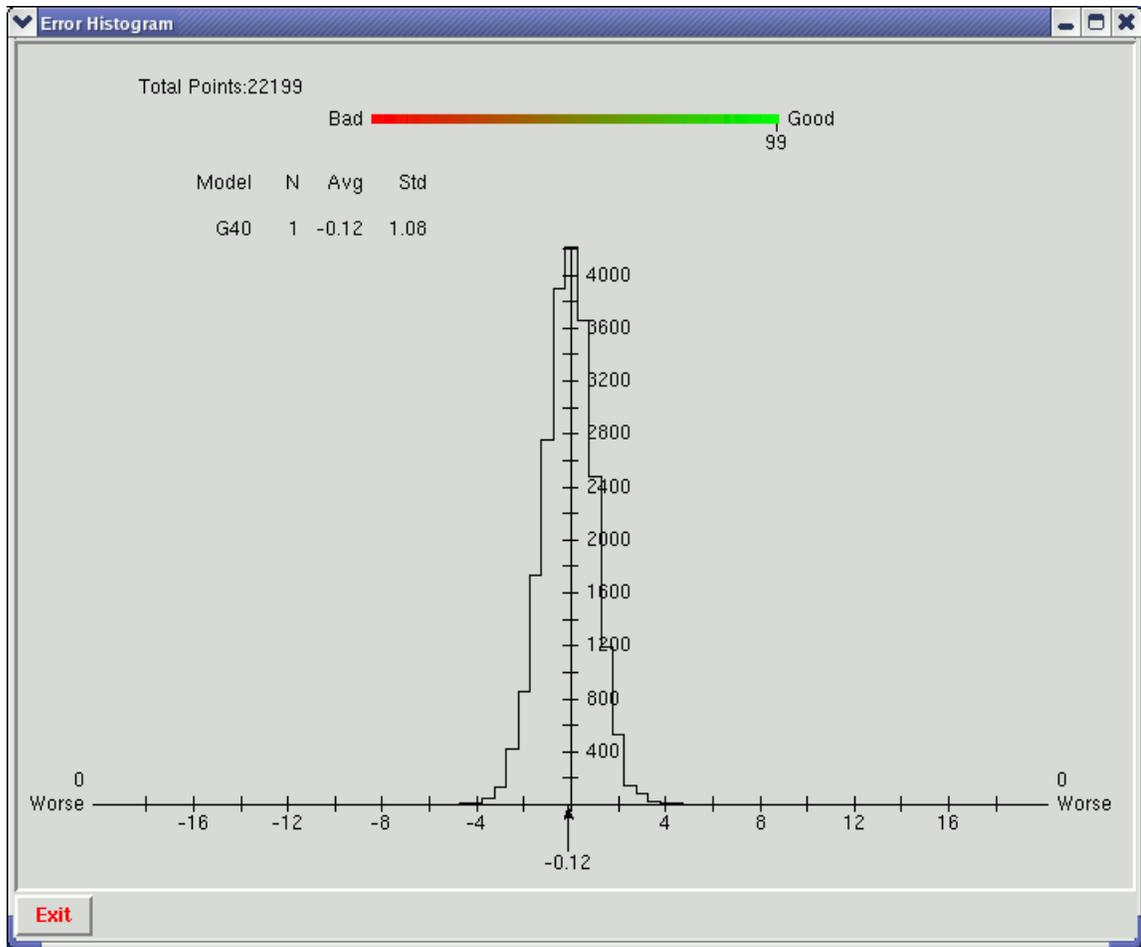


Fig 7. Error histogram after applying the correction grid from only the single best GFS40 first period MaxT analog in the 40 days prior to July 13 (that dated July 9, 2008). This histogram scored 98.82.

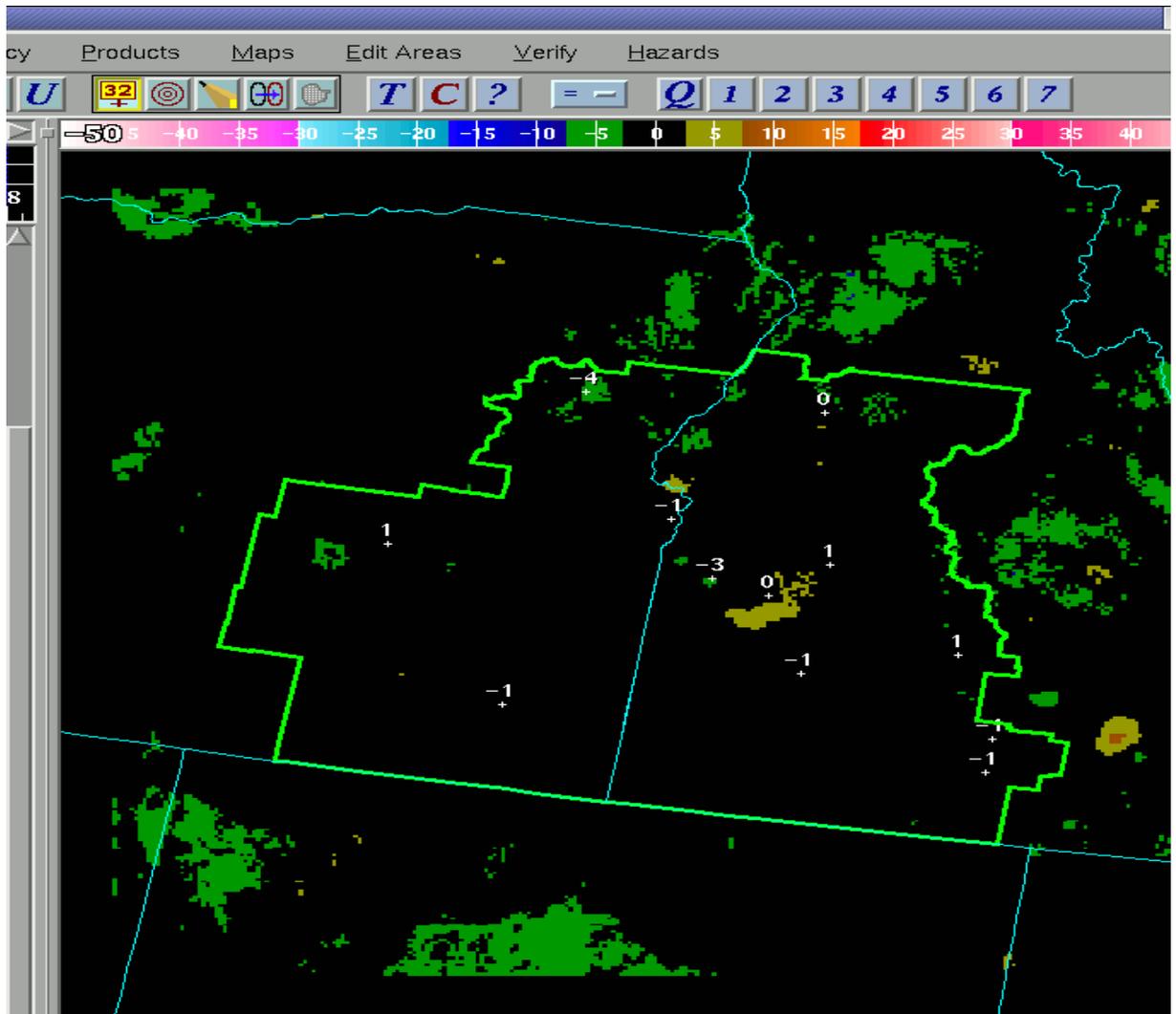


Fig 8. Errors after applying the correction grid made by the GFS40 first period MaxT for July 9, 2008 to the July 13 GFS40 first period MaxT. This one correction successfully removed nearly every error greater than ± 3 F degrees.

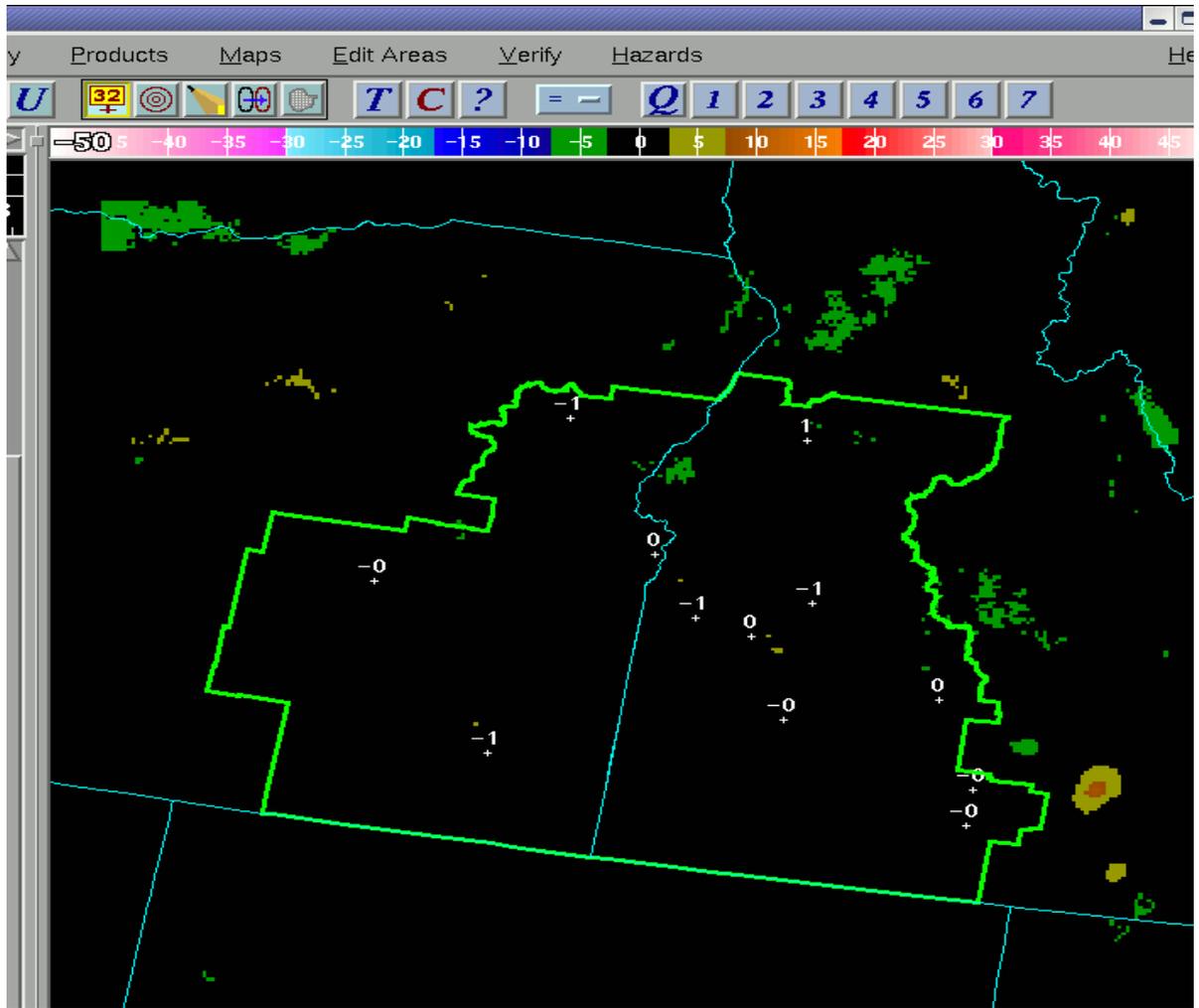


Fig 9. MaxT errors after applying the 40/12 analog correction to the original first period NAM12 for July 13, 2008. This correction had a histogram score of 99.38 within the BOICWA.

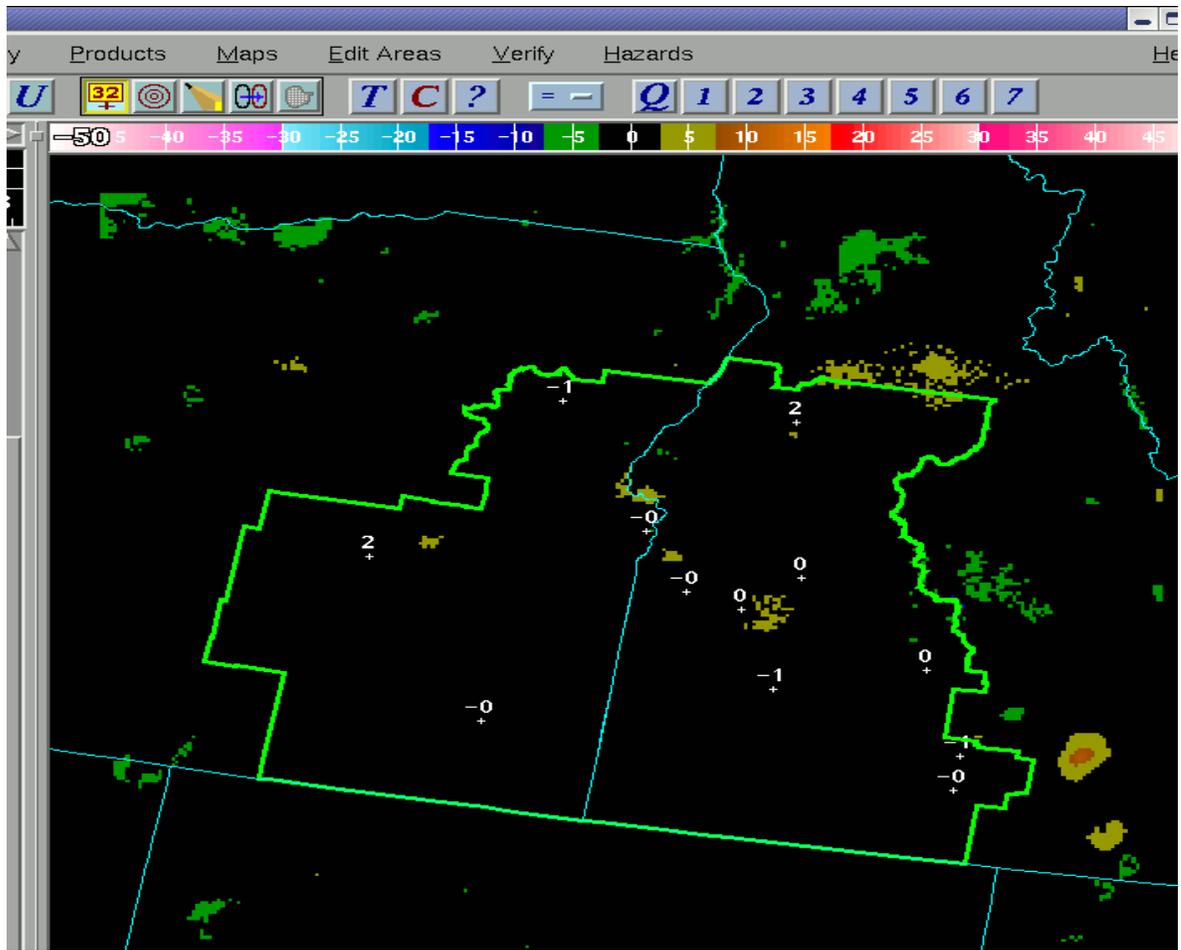


Fig 10. Same as Fig 9 but for 1st period ADJMAV. The histogram for this model scored 99.03.

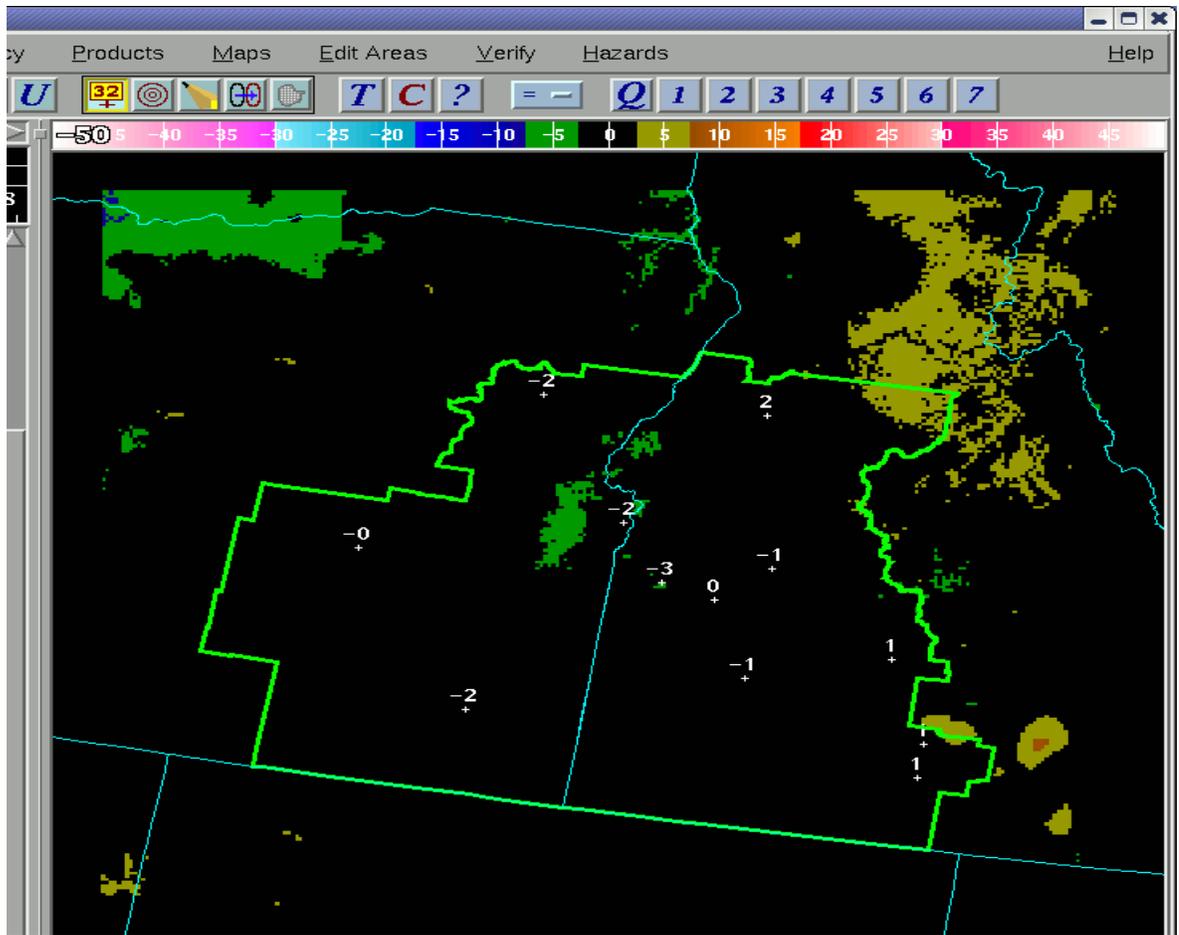


Fig. 11. Same as Fig 9 but for 1st period ADJMET. The MaxT histogram for this model scored 98.70.

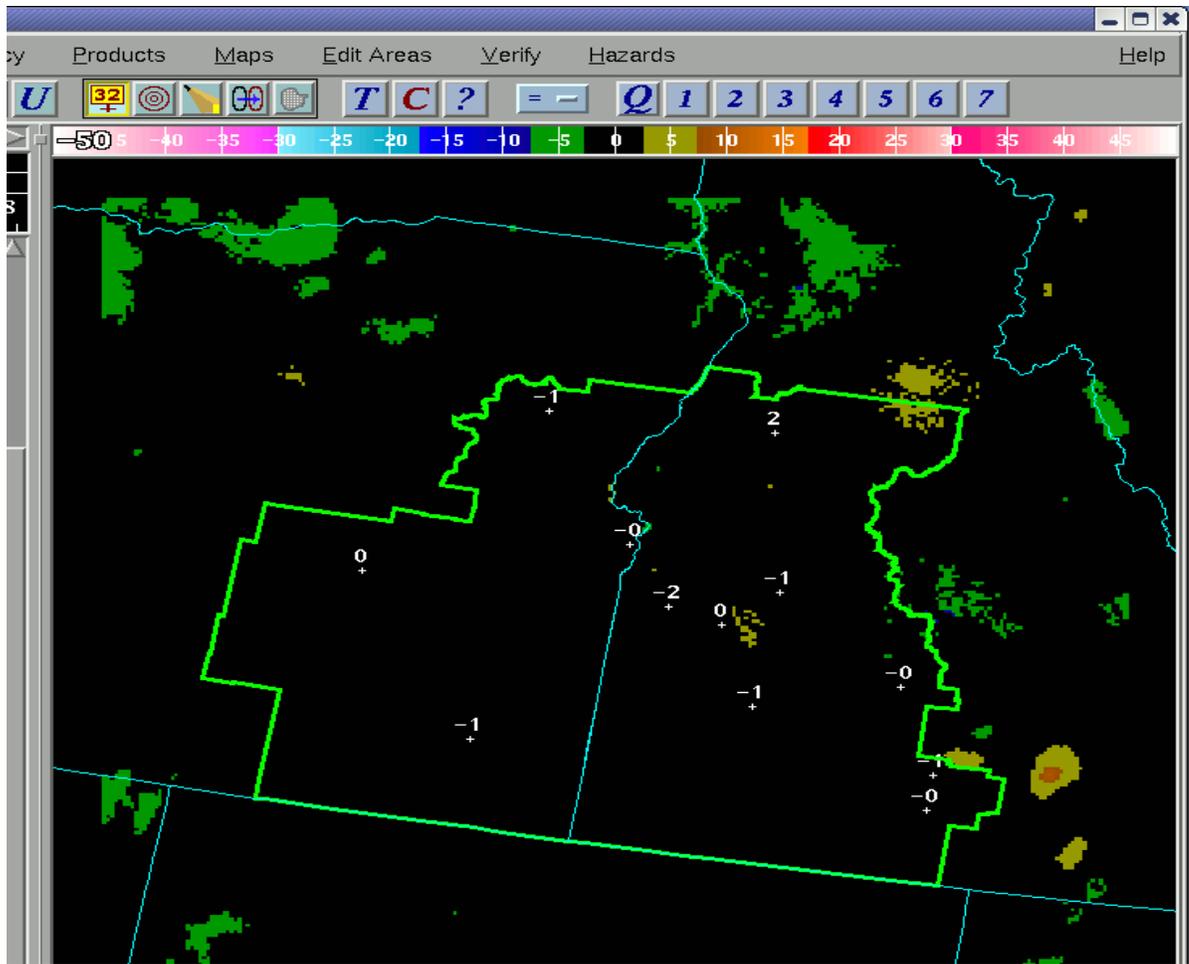


Fig 12. Same as Fig 9 but for 1st period MosGuide. The MaxT histogram for this model scored 99.16.

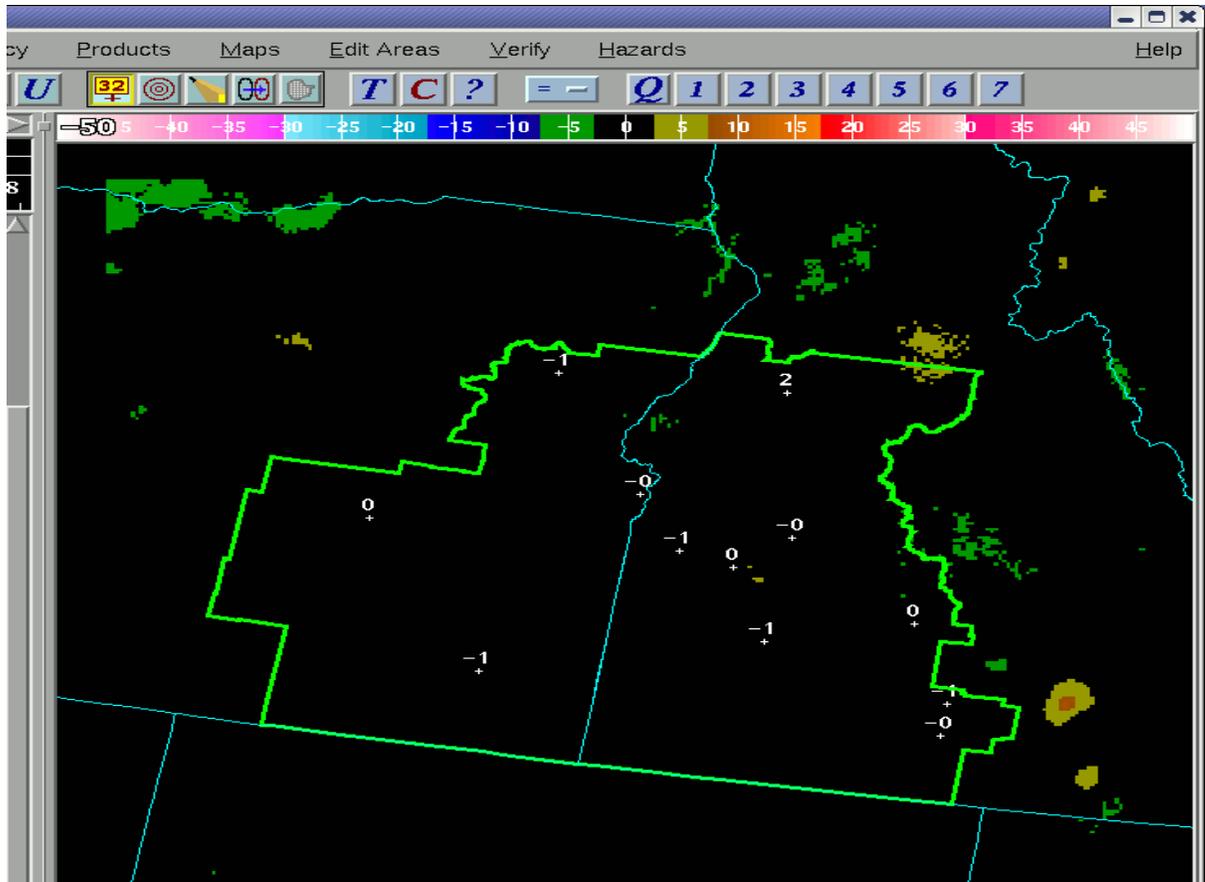


Fig 13. Same as Fig 9 but for all five 40/12 Analog-corrected models (GFS40, NAM12, ADJMAV, ADJMET, and MosGuide) blended together. The blend has successfully removed nearly every remaining error greater than $\pm 3F$ degrees from Figs 6, 9, 10, 11, and 12. This shows that the various models in the blend were correcting errors made by each other.

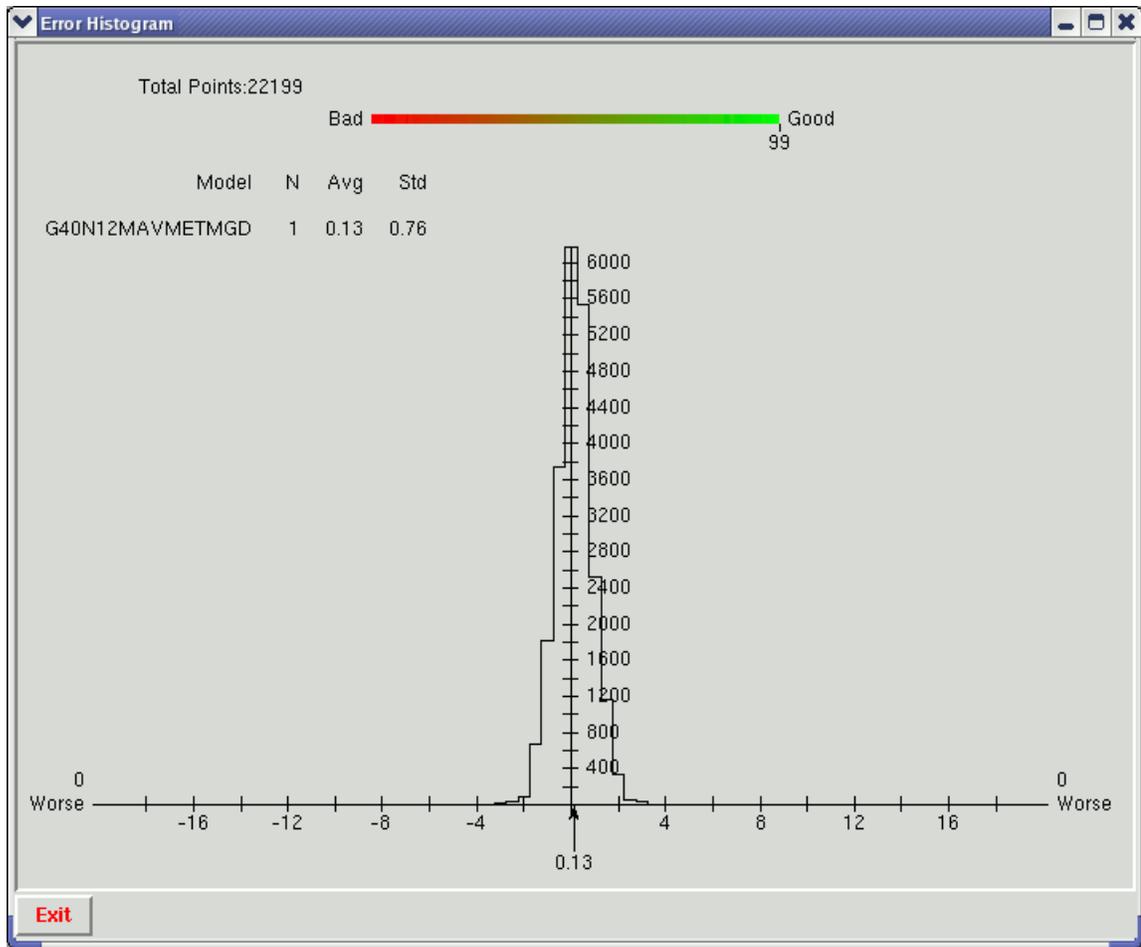


Fig 14. MaxT error histogram for the blend in Fig 13. The 99.40 score was better than for any individual model in the blend and stands as the highest histogram score ever produced for a temperature forecast at WFO BOI on the BOI CWA.

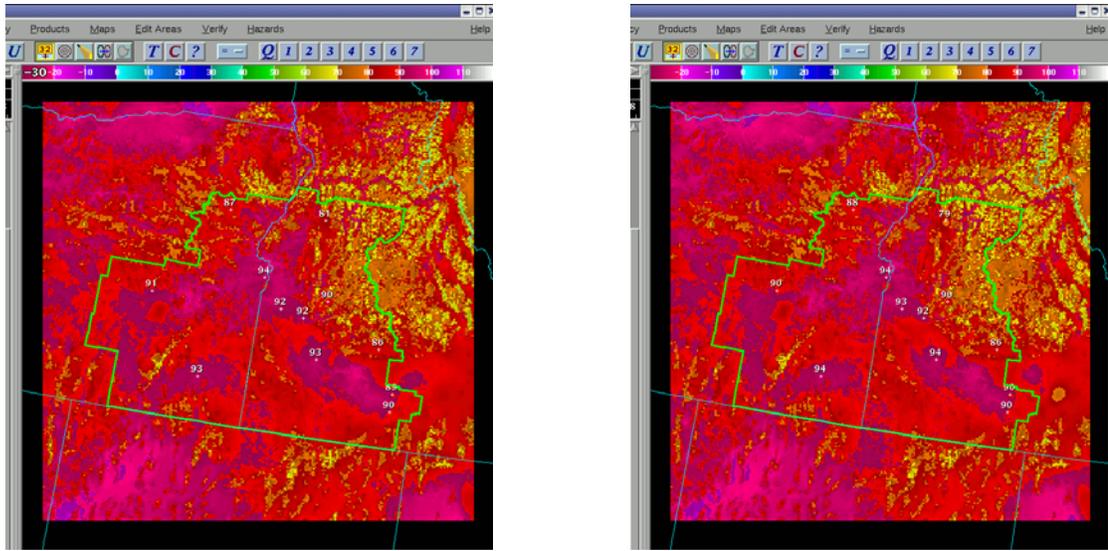


Fig 15. 40/12 1st period MaxT Analog blend (left) and observed MaxT (right) for July 13, 2008.